REMARKS

The examiner objected to the drawings in Applicant's previous amendment.

Applicant believes that the drawings previously submitted fully comply with the objections raised in the prior and the current office actions. No amendment to the drawings (FIG. 8A "104") is needed because 104 is being used to designate two different occurrences of tactile sensors 104. However, Applicant has amended the specification to correct 104 to 107, as mentioned below.

Specification

Applicant has amended the specification to address the informalities mentioned by the examiner and indicated the correct places to insert the amended paragraphs. Enclosed is a substitute specification with these changes. No new matter has been added.

35 U.S.C § 103

The examiner rejected claims 1, 12, 14, 18 and 20-23 under 35 U.S.C. 103(a), as being unpatentable over Choy et al. (US 6695770) in view of Yee et al. (US 6016385).

The examiner stated:

> **7. In regards to claims 1 and 14, Choy discloses a virtual reality encounter system comprising: a mannequin coupled to a computer system wherein the mannequin is fitted with appropriate sensors that are connected to the computer system to transmit to another location and user device over a network (3:23-25), a headset, to display morphing animations and animated textures on the appropriate avatar (9:65-10:6) and a processor that overlays a virtual environment over one or more portions of the video image to form a virtual scene (8:47-58 and 9:65-10:6), Choy lacks explicitly stating the use of a camera supported by the mannequin.**
>
> **8. In related prior art, Yee discloses a rohot system wherein an operator controls the robot and receives sensory information from the robot, including a pair of cameras corresponding to the remote user's eyes coupled to the rohot for receiving a video image where the cameras send the video images via a communication network to the user (5:11-37). One skilled in the art would recognize the advantages of providing video signals to a remote user.**
>
> **9. Therefore it would have been ohvious to one skilled in the art at the time to combine the camera configuration of Yee with the two person configuration of Choy to provide a more realistic experience to hoth users.**

Claim 1 calls for "a virtual reality encounter system includes a mannequin, a camera supported by the mannequin, the camera for capturing an image of a scene; a processor receiving the image from the camera, overlaying a virtual environment over one or more portions of the image to form an image of a virtual scene and sending the image of the virtual scene to a communications network; and a set of goggles…."

The examiner notes that: "**Choy lacks explicitly stating the use of a camera supported by the mannequin.**" However, the examiner argues that Choy discloses: "**a processor that overlays a virtual environment over one or more portions of the video image to form a virtual scene (8:47-58 and 9:65-10:6),**"

Applicant contends that examiner's comments are not directed to the claimed subject matter. Applicant respectfully wishes to emphasize that claim 1 states "a processor receiving the image from the camera." If the examiner admits that Choy does not describe or suggest "the use of a camera supported by the mannequin", it does not logically follow that Choy discloses the subsequent feature which requires that "a processor receiving the image from the camera, overlaying a virtual environment over one or more portions of the image to form an image of a virtual scene and sending the image of the virtual scene to a communications network."

Furthermore, the discussion in Choy pointed out by the examiner in both the office action, and the "response to arguments" seems to validate Applicant's position. For example, Choy at col. 8, lines 53-58 states:

> **To provide a database of images photographic or video recording is mode of a variety of scenes (sex or otherwise) each with a blue background so that this can be superimposed on selected backgrounds such as landscape. Frame by Frame processing is then conducted to create library of sex positions.**

As such, because the camera is not described as supported by the mannequin in Choy, Choy explicitly teaches building a database of images rather than using pictures captured by a camera in order to perform overlaying of images from the database with selected backgrounds desired by the user. Applicant's claim 1 is distinct over Choy.

Yee does not cure the foregoing deficiencies of Choy. The examiner states that: "**In related prior art, Yee discloses a robot system wherein an operator controls the robot and receives sensory information from the robot, including a pair of cameras corresponding to the remote user's eyes coupled to the robot for receiving a video image where the cameras send the video images via a communication network to the user (5:11-37). One skilled in the art**

would recognize the advantages of providing video signals to a remote user." However, Yee neither describes nor suggests the processor receiving the image from the camera, overlaying a virtual environment over one or more portions of the image to form an image of a virtual scene and sending the image of the virtual scene to a communications network. The passage at (5:11-37) in Yee pointed by the examiner does not suggest this feature. In contrast, Yee describes a tele-operated intelligent robot system in which the robot generates sensory signals in response to its interaction with the environment and relates such signals to actuators that give the operator the same sensory feelings[1]. Video information from each camera in Yee is relayed through the command system 14 to respective screens on a helmet. At col. 3, line 67 to col. 4, line 5, Yee states:

> **The antiphon control unit 16 gathers information concerning the robot environment, and a communication system 14 which transmits the commands from command station 12 to the antiphon 16 and transmits back to the command control unit information related to the conditions at the antiphon site.**

It appears that communication system 14 in Yee merely transmits signal back and forth between control unit 16 and command station 12. However, taken the entire document into account, Yee does not describe " a processor receiving the image from the camera", let alone "overlaying a virtual environment over one or more portions of the image to form an image of a virtual scene and sending the image of the virtual scene to a communications network." Therefore, Yee does not cure the deficiencies in Choy and the alleged combination of Choy and Yee neither describes nor suggests claim 1.

Claim 14 is allowable over Choy and Yee at least for the reasons discussed in claim 1.

Claims 12 and 22 depend directly from claim 1 and 14, respectively. Claim 12 recites that "the set of goggles, comprises a receiver to receive the virtual scene." As presented above, Choy fails to describe that "a processor receiving the image from the camera, overlaying a virtual environment over one or more portions of the image to form an image of a virtual scene and sending the image of the virtual scene to a communications network," which is an indispensable

[1] See col. 3, Summary section in Yee.

prior step to form a virtual scene for claim 12. Therefore, claims 12 and 22 are allowable over Choy.

Claims 18, 20, 21, and 23 are allowable over Choy and Yee for at least the reasons given in claims 1 and 14.

The examiner rejected claims 2-6, 10, 11, 13 and 15-17 under 35 U.S.C. 103(a), as being unpatentable over Choy in view of Yee as applied to claim 1 above, and further in view of Dundon (US 7046151).

Applicant contends that claims 2-6, 10, 11, 13 and 15-17 17 are allowable over Choy in view of Yee and further in view of Dundon at least for the reasons discussed in claim 1, in that Dundon does not cure the deficiencies of the combination of referenced applied to claim 1.

The examiner rejected claims 7, 8, 9 rejected under 35 U.S.C. 103(a), as being unpatentable over Choy in view of Yee and Dundon as applied to claim 6 above, and further in view of Abbasi (US 6786863).

Applicant contends that claims 7, 8, and 9 are allowable over Choy in view of Yee and further in view of Dundon and further in view of Abbasi at least for the reasons discussed in claim 1, in that Abbasi does not cure the deficiencies of the combination of references.

While Abbasi mentions: "Once the compressed video arrives at the second computing device, it is presented on a graphic display. This provides a visual perception of the contact episode embodied in the manipulation of the mechanical surrogates." Abbasi does not teach a set of goggles. Rather, Abbasi teaches a display attached to the computer as depicted in FIG. 1. No use would be provided by substituting the display for the set of goggles alleged to be disclosed by Choy.

The examiner rejected claim 19 under 35 U.S.C. 103(a), as being unpatentable over Choy in view of Yee as applied to claim 18 above, and further in view of Abbasi.

Claim 19 is allowable over Choy in view of Yee and further in view of Abbasi at least for the reasons discussed in claim 1, in that Abbasi does not cure the deficiencies of the combination of referenced applied to claim 1.

Applicant has added new claim 24 that further distinguishes over the references. Claim 24 includes the features of "... a first mannequin including ... a first camera supported by the first mannequin ... a second mannequin including ... a second camera supported by the second mannequin, ... a first body suit having motion sensors disposed over the first body suit, ... motion actuators ... and a processor receiving and processing the first image and the second image over a communications network. Claim 24 also includes a set of goggles having a display ... rendering on the display at least one of the first image and the second image ... and a second body suit ... ."

Newly added claim 24 is allowable at least for the reasons discussed above because no combination of Choy, Yee and Abbasi suggest at least a camera supported by both the first and the second mannequins. Claim 24 is also allowable because no combination of the references suggests the arrangement of the first and second mannequins. Claim 25 is allowable with claim 24 as well as for analogous reasons discussed above.

It is believed that all the rejections and/or objections raised by the examiner have been addressed.

In view of the foregoing, applicant respectfully submits that the application is in condition for allowance and such action is respectfully requested at the examiner's earliest convenience.

All of the dependent claims are patentable for at least the reasons for which the claims on which they depend are patentable.

Canceled claims, if any, have been canceled without prejudice or disclaimer.

Any circumstance in which the applicant has (a) addressed certain comments of the examiner does not mean that the applicant concedes other comments of the examiner, (b) made arguments for the patentability of some claims does not mean that there are not other good

reasons for patentability of those claims and other claims, or (c) amended or canceled a claim does not mean that the applicant concedes any of the examiner's positions with respect to that claim or other claims.

Please apply any other charges or credits to deposit account 06-1050.

Respectfully submitted,

Date: 5/15/07

Dennis G. Maloney
Reg. No. 29,670

Fish & Richardson P.C.
225 Franklin Street
Boston, MA 02110
Telephone: (617) 542-5070
Facsimile: (617) 542-8906

21874203.doc

VIRTUAL ENCOUNTERS

TECHNICAL FIELD

This disclosure relates to virtual reality devices, and in particular, using these devices for communication and contact.

BACKGROUND

Two people can be separated by thousands of miles or across a town. With the development of the telephone, two people can hear each other's voice, and, to each of them, the experience is as if the other person was right next to them. Other developments have increased the perception of physical closeness. For example, teleconferencing and Internet cameras allow two people to see each other as well as hear each other over long distances.

SUMMARY

In one aspect, the invention is a virtual encounter system that includes a mannequin coupled to a camera for receiving a video image. The camera sends the video image to a communications network. The virtual encounter system also includes a processor for overlaying a virtual environment over one or more portions of the video image to form a virtual scene and a set of goggles to render the virtual scene.

In another aspect, the invention is a method of having a virtual encounter. The method includes receiving a video image at a camera coupled to a mannequin. The camera sends the video image to a communications network. The method also includes overlaying a virtual environment over one or more

portions of the video image to form a virtual scene and rendering the virtual scene using a set of goggles.

One or more of the aspects above have one or more of the following advantages. The virtual encounter system adds a higher level of perception that two people are in the same place. Aspects of the system allow two people to touch and to feel each other as well as manipulate objects in each other's environment. Thus, a business person can shake a client's hand from across an ocean. Parents on business trips can read to their children at home and put them to bed. People using the system while in two different locations can interact with each other in a virtual environment of their own selection, e.g., a beach or a mountaintop. People can change their physical appearance in the virtual environment so that they seem taller or thinner to the other person or become any entity of their own choosing.

DESCRIPTION OF THE DRAWINGS

FIG. 1 is a view of a virtual encounter system.

FIG. 2A is a view of a left side of a head of a mannequin.

FIG. 2B is a view of a right side of the head of the mannequin.

FIG. 3 is a view of a set of virtual glasses.

FIG. 4 is a view of a wireless earphone.

FIG. 5 is a functional diagram of the virtual encounter system.

FIG. 6 is a signal flow diagram of the virtual encounter system.

FIG. 7A is a view of a user with motion sensors.

FIG. 7B is a view of a robot with motion actuators.
FIG. 8A is a view of a left hand of the robot.
FIG. 8B is a view a left glove worn by the user.
FIG. 9A is a view of a robot with tactile actuators.
FIG. 9B is a view of the user with tactile sensors.
FIG. 10A is a view of a scene with the user in a room.
FIG. 10B is a view of the scene with the user on a beach.
FIG. 11A is a view of an image of the user.
FIG. 11B is a view of a morphed image of the user.

DESCRIPTION

Referring to FIG. 1, a virtual encounter system 10 includes in a first location A, a mannequin 12a, a communication gateway 16a, a set of goggles 20a worn by a user 22a, and two wireless earphones (earphone 24a and earphone 26a) also worn by user 22a. System 10 can further include in a location B, a mannequin 12b, a communication gateway 16b, a set of goggles 20b worn by a user 22b, and two wireless earphones (earphone 24b and earphone 26b) also worn by user 22b. Gateway 16a and gateway 16b are connected by a network 24 (e.g., the Internet).

As will be explained below, when user 22a interacts with mannequin 12a in location A by seeing and hearing the mannequin, user 22a perceives seeing user 22b and hearing user 22b in location B. Likewise, user 22b listens and sees mannequin 12b but perceives listening and seeing user 22a in location A. Details of the gateways 16a and 16b are discussed below. Suffice it to say that the gateways 16a and 16b execute processes to process and transport raw data produced for instance when users 22a and 22b interact with respective mannequins 12a and 12b.

Referring to FIGS. 2A and 2B, each mannequin 12a-12b includes a camera (e.g., camera 30a and camera 30b) positioned in a left eye socket (e.g., left eye socket 34a and left eye socket 34b), and a camera (e.g., camera 36a and camera 36b) positioned in a right eye socket (e.g., right eye socket 38a and right eye socket 38b).

Each mannequin 12a-12b also includes a microphone (e.g., microphone 42a and microphone 42b) positioned within a left ear (e.g., left ear 46a and left ear 46b), and a microphone (e.g., microphone 48a and microphone 48b) positioned within a right ear (e.g., right ear 52a and right ear 52b).

Each mannequin 12a-12b further includes a transmitter (e.g., transmitter 72a and transmitter 72b) containing a battery (not shown). Transmitters 72a-72b send the audio and video signals from the cameras and the microphones to communication gateway 16a-16b.

Referring to FIG. 3, each set of goggles 20a and 20b includes one left display (left display 56a and left display 56b) and one right display (right display 60a and right display 60b). Each set of goggles 20a and 20b includes a receiver (e.g., receiver 70a and receiver 70b) containing a battery source (not shown). Receivers 70a-70b receive the audio and video signals transmitted from processors 16a-16b.

Referring to FIG. 4, each earphone 24a, 24b, 26a and 26b includes a receiver 74 for receiving audio signals from a corresponding microphone 42a, 42b, 48a and 48b an amplifier 75 for amplifying the audio signal and a transducer 76 for broadcasting audio signals.

Referring to FIG. 5, each communication gateway 16a-16b includes an adapter 78a-78b, a processor 80a-80b, memory 84a-84b, an interface 88a-88b and a storage medium 92a-92b (e.g.,

a hard disk). Each adapter 78a-78b establishes a bi-directional signal connection with network 24.

Each interface 88a-88b receives, via transmitter 72a-78b in mannequin 12a-12b, video signals from cameras 30a-30b, 36a-36b and audio signals from microphones 42a-42b, 48a-48b. Each interface 88a-88b sends video signals to displays 56a, 56b in goggles 20a-20b via receiver 70a-70b. Each interface 88a sends audio signals to earphones 24a-24b, 26a-26b in goggles 20a-20b via receiver 74a-74b.

Each storage medium 92a-92b stores an operating system 96a-96b, data 98a-98b for establishing communications links with other communication gateways, and computer instructions 94a-94b which are executed by processor 80a-80b in respective memories 84a-84b to coordinate, send and receive audio, visual and other sensory signals to and from network 24.

Signals within system 10 are sent using a standard streaming connection using time-stamped packets or a stream of bits over a continuous connection. Other examples, include using a direct connection such as an integrated services digital network (ISDN).

Referring to FIG. 6, in operation, camera 30b and camera 36b record video images from Location B. The video images are transmitted wirelessly to communication gateway 16b as video signals. Communication gateway 16b sends the video signals through network 28 to communication gateway 16a. Communication gateway 16b transmits the video signals wirelessly to set of goggles 20a. The video images recorded by camera 30b are rendered on to display 56a, and the video images recorded on camera 36b are rendered on to display 60a.

Likewise, communication gateway 16a and communication gateway 16b work in the opposite direction through network 24,

so that the video images, from location A, recorded by camera 30a are rendered on to display 56b. The video images, recorded by camera 36a are rendered on display 60b.

The sounds received by microphone 42a in location A, are transmitted to earphone 24b and sounds received in location A by microphone 52a are transmitted to earphone 26b. The sounds received by microphone 42b in location B, are transmitted to earphone 24a and sounds received in location B by microphone 52b are transmitted to earphone 26a.

Using system 10, two people can have a conversation where each of the persons perceives that the other is in the same location as them.

Referring to FIGS. 7A and 7B, the user 22a is shown wearing motion sensors 101, over portions of their bodies, and in particular over those portions of the body that exhibit movement. In addition, the mannequins are replaced by robots. For example, a robot 12b' includes a series of motion actuators 103. Each motion actuator 103 placement corresponds to a motion sensor 101 on the user 22a so that each motion sensor activates a motion actuator in the robot that makes the corresponding movement.

For example, when the user 22a moves their right hand, a sensor in the right hand sends a signal through the network to a motion actuator on the robot 12b'. The robot 12b' in turn moves its right hand.

In another example, a user 22a can walk towards a robot 12a' in location A. All the sensors on the user 22a send a corresponding signal to the actuators on the robot 12b' in location B. The robot 12b' in location B performs the same walking movement. The user 22b in location B is not looking in location B but rather through the eyes of the robot 12a' in

location A so that user 22b does see the user 22a in location A walking towards them, but not because the robot 12b' in location B is walking. However, the fact that the robot 12b' in location B is walking enables two things to happen. First, since the user 22a in location A is seeing through the eyes of the robot 12b' in location B and since the robot 12b' in location B is walking enables the user 22a in location A to see what he would see if he were indeed walking in location B. Second, it enables the robot 12b' in location B to meet up with the user 22b in location B.

Referring to FIGS. 8A and 8B, in still other embodiments, tactile sensors 104 are placed on the exterior of a robot hand 102 located in Location A. Corresponding tactile actuators 106 are sewn into an interior of a glove 107 worn by a user in location B. Using system 10, a user in location B can feel objects in Location A. For example, a user can see a vase within a room, walk over to the vase, and pick-up the vase. The tactile sensors-actuators are sensitive enough so that the user can feel the texture of the vase.

Referring to FIGS. 9A and 9B, in other embodiments, sensors are placed over various parts of a robot. Corresponding actuators can be sewn in the interior of a body suit that is worn by a user. The sensors and their corresponding actuators are calibrated so that more sensitive regions of a human are calibrated with a higher degree of sensitivity.

Referring to FIGS. 10A and 10B in other embodiments, user 22a can receive an image of a user 22b but the actual background behind user 22b is altered. For example, user 22b is in a room 202 but user 22a perceives user 22b on a beach 206 or on a mountaintop (not shown). Using conventional video

image editing techniques, the communication gateway 16a processes the signals received from Location B and removes or blanks-out the video image except for the portion that has the user 22b. For the blanked out areas on the image, the communication gateway 16a overlays a replacement background, e.g., virtual environment to have the user 22b appear to user 22a in a different environment. Generally, the system can be configured so that either user 22a or user 22b can control how the user 22b is perceived by the user 22a. Communication gateway 16a using conventional techniques can supplement the audio signals received with stored virtual sounds. For example, waves are added to a beach scene, or eagles screaming are added to a mountaintop scene.

In addition, gateway 16a can also supplement tactile sensations with stored virtual tactile sensations. For example, a user can feel the sand on her feet in the beach scene or a cold breeze on her cheeks in a mountain top scene.

In this embodiment, each storage medium 92a-92b stores data 98a-98b for generating a virtual environment including virtual visual images, virtual audio signals, and virtual tactile signals. Computer instructions 94a-94b, which are executed by processor 80a-80b out of memory 84a-84b, combine the visual, audio, and tactile signals received with the stored virtual visual, virtual audio and virtual tactile signals in data 98a-98b.

Referring to FIGS. 11A and 11B, in other embodiments, a user 22a can receive a morphed image 304 of user 22b. For example, an image 302 of user 22b is transmitted through network 24 to communications gateway 16a. User 22b has brown hair, brown eyes and a large nose. Communications gateway 16a again using conventional imaging morphing techniques alters

the image of user 22b so that user 22b has blond hair, blue eyes and a small noise and sends that image to goggles 20a to be rendered.

Communication gateway 16a also changes the sound user 22b makes as perceived by user 22a. For example, user 22b has a high-pitched squeaky voice. Communication gateway 22b using conventional techniques can alter the audio signal representing the voice of user 22b to be a low deep voice.

In addition, communication gateway 16a can alter the tactile sensation. For example, user 22b has cold, dry and scaling skin. Communications gateway 16a can alter the perception of user 22a by sending tactile signals that make the skin of user 22b seem smooth and soft.

In this embodiment, each storage medium 92a-92b stores data 98a-98b for generating a morph personality. Computer instructions 94a-94b, which are executed by processor 80a-80b out of memory 84a-84b, combine the visual, audio, and tactile signals received with the stored virtual visual, virtual audio and virtual tactile signals of a personality in data 98a-98b.

Thus using system 10 anyone can assume any other identity if it is stored in data 98a-98b.

In other embodiments, earphones are connected to the goggles. The goggles and the earphones are hooked by a cable to a port (not shown) on the communication gateway.

Other embodiments not described herein are also within the scope of the following claims.